EP376588

Abstract:

Abstract not available for EP0376588 Data supplied from the esp@cenet database - Worldwide

----
Courtesy of http://v3.espacenet.com

EUROPEAN PATENT APPLICATION

(71) Applicant: AMERICAN TELEPHONE AND
TELEGRAPH COMPANY
550 Madison Avenue
New York, NY 10022(US)

(72) Inventor: Erving, Richard Henry
165 Riddle Avenue, Unit 6, Long Branch
Monmouth County, New Jersey 07740(US)

(74) Representative: Buckley, Christopher Simon
Thirsk et al
AT&T (UK) LTD. AT&T Intellectual Property
Division 5 Mornington Road
Woodford Green, Essex IG8 OTU(GB)

(54) Computer controlled speakerphone for adapting to a communication line.

(57) A computer controlled speakerphone includes a line adapting arrangement (110, 200, 300) for developing information about an interface between the speakerphone and a hybrid (610) in a communication line and for providing optimal performance during operation by adapting operating parameters of the speakerphone to the line. During a conversation, the line adapting arrangement measures and averages the degree of hybrid reflection that is presented to the speakerphone. This hybrid reflection provides a measure of both the hybrid and a far-end acoustic return. By determining the degree of hybrid reflection, the switching threshold level of the speakerphone for switching between the transmit state and the receive state may be adjusted. Once the expected level of receive speech due to hybrid reflection is known, additional receive speech due to the far-end talker may be accurately determined and the state of the speakerphone switched accordingly. The amount of switched loss required in the transmit and receive speech paths of the speakerphone to maintain stability also may be adjusted by the line adapting arrangement in accordance with the degree of hybrid reflection. By lowering the amount of switched loss, speakerphone switching operation becomes more transparent and can even approach full-duplex for fully digital connections.

EP 0 376 588 A2

Computer Controlled Speakerphone for Adapting to a Communication Line

Background of the Invention

1. Technical Field

5

This invention relates to audio systems and, more particularly, to voice switching circuits which connect to a communication line for providing two-way voice switched communications.

10   2. Description of the Prior Art

The use of analog speakerphones have been the primary hands free means of communicating during a telephone conversation for a great number of years. This convenient service has been obtained at the price of some limitations, however. These speakerphone usually require careful and expensive calibration in order
15   to operate in an acceptable manner. They are also designed to operate in a worst-case electrical environment thereby sacrificing the improved performance that is possible in a better environment.

The operation of conventional analog speakerphones is well known and is described in an article by A. Busala, "Fundamental Considerations in the Design of a Voice-Switched Speakerphone," Bell System Technical Journal, Vol. 39, No. 2, March 1960, pp 265-294. Analog speakerphones generally use a
20   switched-loss technique through which the energy of the voice signals in both a transmit and a receive direction are sensed and a switching decision made based upon that information. The voice signal having the highest energy level in a first direction will be given a clear talking path and the voice signal in the opposite direction will be attenuated by having loss switched into its talking path. If voice signals are not present in either the transmit direction or the receive direction, the speakerphone goes to an "at rest" mode
25   which provides the clear talking path to voice signals in a receive direction favoring speech from a distance speaker. In some modem analog speakerphones, if voice signals are not present in either the transmit direction or the receive direction, the speakerphone goes to an idle mode where the loss in each direction is set to a mid-range level to allow the direction wherein voice signals first appear to quickly obtain the clear talking path.

30   Most high-end analog speakerphones also have a noise-guard circuit to adjust the switching levels according to the level of background noise present. Switching speed is limited by a worst-case time constant that assures that any speech energy in the room has time to dissipate. This limitation is necessary to prevent "self switching", a condition where room echoes are falsely detected as near-end speech.

A disadvantage associated with the conventional analog speakerphone is that it is unable to relate to the
35   hybrid it faces when attached to a telephone line. Even a digital telephone within a private branch exchange (PBX), which does not employ a hybrid, faces an unpredictable hybrid on calls which are routed outside of the PBX. A worst case trans-hybrid loss must therefore be assumed. This assumption requires the insertion of more switched loss than otherwise might he necessary in order to assure that the system will remain stable. A high "break in" threshold is similarly required in order to prevent a bad hybrid from reflecting
40   enough transmit speech to falsely switch the speakerphone into the receive state. Thus the optimal performance possible with this speakerphone in its interface with a hybrid is not achieved.

Summary of the Invention

45

In accordance with the present invention, a speakerphone develops information about an interface between the speakerphone and a hybrid in a communication line for adapting operating parameters of the speakerphone to the line for providing optimal performance during operation.

During a conversation, a control unit such as a computer in the speakerphone measures and averages
50   the degree of hybrid reflection that is presented to the speakerphone. This hybrid reflection or hybrid average provides a measure of both the hybrid and a far-end acoustic return. The hybrid average is determined in the speakerphone through a process whereby a transmit signal is subtracted from a receive signal and the results averaged in a manner that favors the maximum difference between these signals. The received signal is that signal provided to the speakerphone by the hybrid on a receive line and the transmit signal is that signal provided to the hybrid by the speakerphone on a transmit line.

Calculation of the hybrid average provides enhanced operation for the speakerphone in a number of ways. By developing an estimate of the hybrid average, the amount of switched loss required in the transmit and receive speech paths of the speakerphone to maintain stability may be raised or lowered by the computer as appropriate. By lowering the amount of switched loss, speakerphone switching operation becomes more transparent and can even approach full-duplex for fully digital connections.

The estimate of the hybrid average is also used by the computer to determine the switching threshold level of the speakerphone for switching between the transmit state and the receive state. Once the hybrid average is developed, it is used to provide an expected level of receive speech due to reflection. Once this expected level of receive speech is known, additional receive speech due to the far-end talker may be accurately determined and the state of the speakerphone switched accordingly.

Certain boundary conditions are employed in developing the hybrid average. To obtain an accurate representation of the line conditions, for example, hybrid averaging is performed only while the speakerphone is in the transmit state. This insures that receive speech on the receive line during a quiet transmit interval cannot be mistaken for a high level of hybrid return. This averaging thus prevents receive speech, that is not great enough to cause the speakerphone to go into the receive state, from distorting the estimated hybrid average.

To ensure stable operation in a system with an adaptive speakerphone in use at both the near-end and the far-end by both parties, the amount that the hybrid average may improve during any given transmit interval may be limited to a predetermined level such as, for example, 5 dB. Thus, in order for the hybrid average to improve further, a transition to receive and then back to transmit must be made. This insures that the far-end speakerphone has also had an opportunity to go into the transmit state and has similarly adapted. Each speakerphone is therefore able to reduce its inserted loss down to a point of balance in a systematic manner. Limiting the amount of change in the hybrid average in this manner also allows the adaptive speakerphone to to be operable with other adaptive speakerphones such as echo-canceling speakerphones that present a varying amount of far-end echo as they adapt.

Brief Description of the Drawing

FIG. 1 is a block representation of the major functional components of a computer controlled adaptive speakerphone operative in accordance with the principles of the invention;

FIG. 2 is a partial schematic of the speakerphone including a calibration circuit, an amplifier for remotely provided speech signals, a microphone and an associated amplifier and multiplexers employed in this invention;

FIG. 3 is a partial schematic of the speakerphone including mute controls and high pass filters employed in this invention;

FIG. 4 is a schematic of a programmable attenuator and a low pass filter employed in a transmit section of this invention;

FIG. 5 is a schematic of a programmable attenuator and a low pass filter employed in a receive section of this invention;

FIG. 6 depicts a general speakerphone circuit and two types of coupling that most affect its operation;

FIG. 7 is a state diagram depicting the three possible states of the speakerphone of FIG. 1;

FIG. 8 depicts a flow chart illustrating the operation of the speakerphone of FIG. 1 in determining whether to remain in an idle state or move from the idle state to a transmit or a receive state;

FIG. 9 depicts a flow chart illustrating the operation of the speakerphone of FIG. 1 in determining whether to remain in the transmit state or move from the transmit state to the receive state or idle state;

FIG. 10 depicts a flow chart illustrating the operation of the speakerphone of FIG. 1 in determining whether to remain in the receive or move from the receive state to the transmit state or idle state;

FIG. 11 are illustrative waveforms which depict impulse and composite characterizations of an acoustic environment performed by the speakerphone of FIG. 1;

FIG. 12 is a block representation of the functional components of a speakerphone operable in providing echo suppression loss insertion;

FIG. 13 depicts a flow chart illustrating the operation of the speakerphone of FIG. 12 in application of echo suppression loss insertion; and

FIG. 14 are waveforms illustrating the application of echo suppression loss insertion.

Detailed Description

FIG. 1 is a functional block representation of a computer controlled adaptive speakerphone 100 operative in accordance with the principles of the invention. As shown, the speakerphone generally comprises a transmit section 200, a receive section 300, and a computer 110. A microcomputer commercially available from Intel Corporation as Part No. 8051 may be used for computer 110 with the proper programming. A microphone 111 couples audio signals to the speakerphone and a speaker 112 receives output audio signals from the speakerphone.

By way of operation through illustration, an audio signal provided by a person speaking into the microphone 111 is coupled into the transmit section 200 to a multiplexer 210. In addition to being able to select the microphone speech signal as an input, the multiplexer 210 may also select calibration tones as its input. These calibration tones are provided by a calibration circuit 113 and are used, in this instance, for calibration of the hardware circuitry in the transmit section 200.

Connected to the multiplexer 210 is a mute control 211 which mutes the transmit path in response to a control signal from the computer 110. A high pass filter 212 connects to the mute control 211 to remove the room and low frequency background noise in the speech signal. The output of the high pass filter 212 is coupled both to a programmable attenuator 213 and to an envelope detector 214. In response to a control signal from the computer 110, the programmable attenuator 213 inserts loss in the speech signal in three and one half dB steps up to a total of sixteen steps, providing 56 dB of total loss. This signal from the programmable attenuator 213 is coupled to a low pass filter 215 which removes any spikes that might have been generated by the switching occurring in the attenuator 213. This filter also provides additional signal shaping to the signal before the signal is transmitted by the speakerphone over audio line 101 to a hybrid (not shown). After passing through the envelope detector 214, the speech signal from the filter 212 is coupled to a logarithmic amplifier 216, which expands the dynamic range of the speakerphone to approximately 60 dB for following the envelope of the speech signal.

The receive section 300 contains speech signal processing circuitry that is functionally the same as that found in the transmit section 200. A speech signal received over an input audio line 102 from the hybrid is coupled into the receive section 300 to the multiplexer 310. Like the multiplexer 210, the multiplexer 310 may also select calibration tones for its input, which are provided by the calibration circuit 113. Connected to the multiplexer 310 is a mute control 311 which mutes the receive path in response to a control signal from the computer 110. A high pass filter 312 is connected to the mute control 311 to remove the low frequency background noise from the speech signal.

The output of the high pass filter 312 is coupled both to an envelope detector 314 and to a programmable attenuator 313. The envelope detector 314 obtains the signal envelope for the speech signal which is then coupled to a logarithmic amplifier 316. This amplifier expands the dynamic range of the speakerphone to approximately 60 dB for following the envelope of the receive speech signal. The programmable attenuator 313, responsive to a control signal from the computer 110, inserts loss in the speech signal in three and one half dB steps in sixteen steps, for 56 dB of loss. This signal from the programmable attenuator 313 is coupled to a low pass filter 315 which removes any spikes that might have been generated by the switching occurring in the attenuator 313. This filter also provides additional signal shaping to the signal before the signal is coupled to the loudspeaker 112 via an amplifier 114.

The signals from both the logarithmic amplifier 216 and the logarithmic amplifier 316 are multiplexed into an eight-bit analog-to-digital converter 115 by a multiplexer 117. The converter 115 presents the computer 110 with digital information about the signal levels every 750 microseconds.

The computer 110 measures the energy of the incoming signals and develops information about the signal and noise levels. Both a transmit signal average and a receive signal average are developed by averaging samples of each signal according to the following equation:

$$\hat{y}_t = \begin{cases} \hat{y}_{t-1} + \dfrac{|s|_t - \hat{y}_{t-1}}{4} & \text{if } |s|_t \geq \hat{y}_{t-1} \\[3mm] \hat{y}_{t-1} + \dfrac{|s|_t - \hat{y}_{t-1}}{32} & \text{if } |s|_t < \hat{y}_{t-1} \end{cases}$$

where
Sampling rate = 1333 per second

4

$|s|_t$ = new sample
$\hat{y}_{t-1}$ = old average
$\hat{y}_t$ = new average

This averaging technique tends to pick out peaks in the signal applied. Since speech tends to have many peaks rather than a constant level, this average favors detecting speech.

Both a transmit noise average and a receive noise average are also developed. The transmit noise average determines the noise level of the operating environment of the speakerphone. The receive noise average measures the noise level on the line from the far-end party. The transmit noise average and the receive noise average are both developed by measuring the lowest level seen by the converter 115. Since background noise is generally constant, the lowest samples provide a reasonable estimate of the noise level. The transmit and receive noise averages are developed using the following equation:

$$\hat{y}_t = \begin{cases} \hat{y}_{t-1} + \dfrac{|s|_t - \hat{y}_{t-1}}{4096} & \text{if } |s|_t \geq \hat{y}_{t-1} \\[4mm] \hat{y}_{t-1} + \dfrac{|s|_t - \hat{y}_{t-1}}{4} & \text{if } |s|_t < \hat{y}_{t-1} \end{cases}$$

where
Sampling rate = 1333 per second
$|s|_t$ = new sample
$\hat{y}_{t-1}$ = old average
$\hat{y}_t$ = new average

This equation strongly favors minimum values of the envelope of the applied signal, yet still provides a path for the resulting average to rise when faced with a noisier environment.

Two other signal levels are developed to keep track of the switching response and singing margin of the speakerphone. These signal levels are the speech level that is present after being attenuated by the transmit attenuator 213 and the speech level that is present after being attenuated by the receive attenuator 313. In the speakerphone, these two levels are inherently known due to the fact that the computer 110 directly controls the loss in the attenuators 213 and 313 in discrete amounts, 3.5 dB steps with a maximum loss of 56 dB in each attenuator. All of these levels are developed to provide the computer 110 with accurate and updated information about what the current state of the speakerphone should be.

As in all speakerphones, the adaptive speakerphone needs to use thresholds to determine its state. Unlike its analog predecessors, however, those thresholds need not be constant. The computer 110 has the ability to recalibrate itself to counteract variation and aging of hardware circuitry in the speakerphone. This is achieved by passing a first and a second computer-generated test tone through the transmit path and the receive path of the hardware circuitry and measuring both responses.

These test tones are generated at a zero dB level and a minus 20 dB level. The difference measured between the zero dB level tone and the minus 20 dB level tone that passes through the speakerphone circuitry is used as a base line for setting up the thresholds in the speakerphone. First, by way of example, the zero dB level tone is applied to the transmit path via multiplexer 210 and that response measured by the computer 110. Then the minus 20 dB tone is similarly applied to the transmit path via multiplexer 210 and its response measured by the computer. The difference between the two responses is used by the computer as a basic constant of proportionality that represents "20 dB" of difference in the transmit path circuitry. This same measurement is similarly performed on the receive path circuitry by applying the two test tones via multiplexer 310 to the receive path. Thus, a constant of proportionality is also obtained for this path. The number measured for the receive path may be different from the number measured by the transmit path due to hardware component variations. The computer simply stores the respective number for the appropriate path with an assigned value of minus 20 dB to each number. Once the computer has determined the number representing minus 20 dB for each path, it is then able to set the required dB threshold levels in each path that are proportionally scaled to that path's number. Also, because of the relative scaling, the common thresholds that are set up in each path always will be essentially equal even though the values of corresponding circuit components in the paths may differ considerably.

As part of the calibration process, the speakerphone also measures the acoustics of the room in which

5

it operates. Through use of the calibration circuit 113, the speakerphone generates a series of eight milliseconds tone bursts throughout the audible frequency of interest and uses these in determining the time-domain acoustic response of the room. each tone burst is sent from the calibration circuit 113 through the receive section 300 and out the loudspeaker 112. The integrated response, which is reflective of the echoes in the room from each tone burst, is picked up by the microphone 111 and coupled via the transmit section 200 to the computer 110 where it is stored as a composite response pattern, shown in FIG. 11 and described in greater detail later herein. This response is characterized by two important factors: the maximum amplitude of the returned signal, and the duration of the echoes. The amplitude of the returned signal determines what level of transmit speech will be required to break in on receive speech. The greater the acoustic return, the higher that threshold must be to protect against self-switching. The duration of the echoes determine how quickly speech energy injected into the room will dissipate, which controls how fast the speakerphone can switch from a receive to a transmit state. If the room acoustics are harsh, therefore, the speakerphone adapts by keeping switching response on a par with that of a typical analog device. But when acoustics are favorable, it speeds past this switching time and lowers break in thresholds to provide a noticeable improvement in performance.

The concept of self-calibration is also applied to the speakerphone's interface to a hybrid. During a conversation, the computer measures the degree of hybrid reflection that it sees. This hybrid reflection provides a measure of both the hybrid and far-end acoustic return. Its average value is determined using the following equation:

$$
\hat{H}_t =
\begin{cases}
\hat{H}_{t-1} + \dfrac{\left(\hat{R}_t - \hat{T}_t\right) - \hat{H}_{t-1}}{4096} & \text{if } \left(\hat{R}_t - \hat{T}_t\right) \geq \hat{H}_{t-1} \\[3mm]
\hat{H}_{t-1} + \dfrac{\left(\hat{R}_t - \hat{T}_t\right) - \hat{H}_{t-1}}{4} & \text{if } \left(\hat{R}_t - \hat{T}_t\right) < \hat{H}_{t-1}
\end{cases}
$$

where
Sampling rate = 1333 per second
$\hat{R}_t$ = receive signal average
$\hat{T}_t$ = transmit signal average
$\hat{H}_{t-1}$ = old hybrid average
$\hat{H}_t$ = new hybrid average

This equation develops the hybrid average value by subtracting a transmit signal from a receive signal and then averaging these signals in a manner that favors the maximum difference between them. The receive signal is that signal provided to the speakerphone by the hybrid on the receive line and the transmit signal is that signal provided to the hybrid by the speakerphone on the transmit line. By developing an estimate of the hybrid average, the amount of switched loss required in the speakerphone to maintain stability may be raised or lowered. By lowering the amount of switched loss, speakerphone switching operation becomes more transparent and can even approach full-duplex for fully digital connections.

The estimate of the hybrid average is also used to determine the switching threshold level of the speakerphone in switching from the transmit state to the receive state (receive break in). Since the estimate of the hybrid average is used to develop an expected level of receive speech due to reflection, additional receive speech due to the far-end talker may be accurately determined and the state of the speakerphone switched accordingly.

To obtain an accurate representation of the line conditions, hybrid averaging is performed only while the speakerphone is in the transmit state. This insures that receive speech on the receive line during a quiet transmit interval cannot be mistaken for a high level of hybrid return. This averaging therefore prevents receive speech, that is not great enough to cause the speakerphone to go into the receive state, from distorting the estimated hybrid average.

Another boundary condition employed in developing this hybrid average is a limitation on the acceptable rate of change of transmit speech. If transmit speech ramps up quickly, then the possibility of sampling errors increases. To avoid this potential source of errors, the hybrid average is only developed during relatively flat intervals of transmit speech (the exact slope is implementation-dependent).

To ensure stable operation with an adaptive speakerphone in use at both the near-end and the far-end by both parties, the amount that the hybrid average may improve during any given transmit interval is also

6

limited. In the adaptive speakerphone 100, for example, the hybrid average is allowed to improve no more than 5 dB during each transmit state. In order for the hybrid average to improve further, a transition to receive and then back to transmit must be made. This insures that the far-end speakerphone has also had an opportunity to go into the transmit state and has similarly adapted. Thus, each speakerphone is able to reduce its inserted loss down to a point of balance in a monotonic fashion. Limiting the amount of change in the hybrid average during a transmit interval also allows this speakerphone to be operable with other adaptive speakerphones such as echo-canceling speakerphones that present a varying amount of far-end echo as they adapt.

For ease of operation and for configuring the speakerphone, a user interface 120 through which the user has control over speakerphone functions is provided internal to the speakerphone 100. This interface includes such speakerphone functions as ON/OFF, MUTE and VOLUME UP/DOWN. The user interface also includes a button or other signaling device for initiating the recalibration process. Should the user relocate his or her speakerphone, pressing this button will perform an acoustic calibration to the new environment. In addition, the recalibration process checks the operational readiness of and recalibrates the internal hardware circuitry, and resets the volume level of the speakerphone to the nominal position.

Referring now to FIGS. 2 and 3, there is shown a partial schematic of the speakerphone 100 including the multiplexers 210 and 310, mute controls 211 and 311, the calibration circuit 113, the microphone 111 and its associated amplifier 117, amplifier 135 for the remotely provided speech signals, and high pass filters 211 and 311.

Shown in greater detail is the microphone 111 which, in this circuit arrangement, is an electret microphone for greater sensitivity. This microphone is AC coupled via a capacitor 116 to an amplifier 117 which includes resistors 118 and 119 for setting the transmit signal gain from the microphone 111. From the amplifier 117, the speech signal is sent to the multiplexer 210 in the transmit section 200.

Also shown in greater detail is the calibration circuit 113 which receives a two-bit input from the computer 110 on lines designated as CALBIT UP and CALBIT DOWN. This two-bit input provides the tone burst signal used in the hardware circuitry and acoustic calibration processes. Three states from the two-bit input are defined and available: LOW reflects a zero level signal where the input signals on both CALBIT UP and CALBIT DOWN are one; HIGH reflects a condition where the input signals to both CALBIT UP and CALBIT DOWN are zero; and MIDDLE reflects a condition where, for example, the CALBIT UP signal is one and the CALBIT DOWN signal is zero. By alternately presenting and removing the respective input signals to both CALBIT UP and CALBIT DOWN in a desired sequence, a tone burst is generated which starts from ground level, goes up to some given positive voltage level, then down to some given negative voltage level, then returns back to ground level.

The CALBIT UP and CALBIT DOWN signals are respectively provided as input signals to an amplifier 121 via a first series connection, comprising diode 122 and resistor 123, and a second series connection, comprising diode 124 and resistor 125. The amplifier 121 and associated circuitry, capacitor 127 and resistor 128, are used to generate the desired output level reflective of the summation of the two input signals. A resistor divider, comprising resistors 156 and 157, provides an offset voltage to the non-inverting input of amplifier 121. Resistor divider, comprising resistors 129 and 130, provide the 20 dB reduction of the signal level from amplifier 121. This reduction is used for the comparison measurement when the speakerphone performs the electrical calibration process. Thus the signal on line 131 is 20 dB less than the signal on line 132. Both of these two signals are coupled to the multiplexers 210 and 310.

A receive audio input level conversion circuit, comprising amplifier 135, resistors 136, 137 and 138, and also capacitor 139, is connected to audio input line 102 for terminating this line in 600 ohms. This signal is coupled from the amplifier 135 to the multiplexer 310 along with the tone signal from amplifier 121 for further processing.

The output of the multiplexer 210 is provided over line 138 to a mute control 211 which mutes the transmit path in response to a control signal from the computer 110 over line 140. Similarly, the output of the multiplexer 310 is provided over line 139 to a mute control 311 which mutes the receive path in response to a control signal from the computer 110 over line 141. Respectively connected to the mute controls 211 and 311 are high pass filters 212 and 213. These high pass filters are essentially identical and are designed to remove the low frequency background noise in the speech signal. Filter 212 comprises a follower amplifier 217, and associated circuitry comprising capacitors 218 and 219, and resistors 220 and 221. The output of filter 212 is coupled over line 142 to the programmable attenuator 213 shown in FIG. 4. And filter 312 comprises a follower amplifier 317, and associated circuitry comprising capacitors 318 and 319, and resistors 320 and 321. The output of filter 312 is coupled over line 143 to the programmable attenuator 313 shown in FIG. 5.

Referring now to FIG. 4, there is shown a detail schematic of the programmable attenuator 213. This

attenuator comprises multiple sections which are formed by passing the output of an amplifier in one section through a switchable voltage divider and then into the input of another amplifier. The signal on line 142 from the high pass filter 212 is coupled directly to a first section of the attenuator 213 comprising a voltage divider consisting of resistors 222 and 223, a switch 224 and a follower amplifier 226. When the

5 switch 224 is closed shorting resistor 222, the voltage developed across the voltage divider essentially will be the original input voltage, all of which develops across resistor 223. Once the switch is opened, in response to a command from the computer 110, the signal developed at the juncture of resistors 222 and 223 is reduced from that of the original input voltage level to the desired lower level. The loss is inserted in each section of the attenuator in this manner.

10   Thus in operation, a speech signal passing through the first section of the attenuator is either passed at the original voltage level or attenuated by 28 dB. If the switch is turned on, i.e., the resistor 222 shorted out, then no loss is inserted. If the switch is turned off, then 28 dB of loss is inserted. The signal then goes through a second similar section which has 14 dB of loss. This second section of the attenuator 213 comprises a voltage divider consisting of resistors 227 and 228, a switch 229 and a follower amplifier 230.

15 This second section is followed by a third section which has 7 dB of loss. This third section of the attenuator 213 comprises a voltage divider consisting of resistors 231 and 232, a switch 233 and a follower amplifier 234. A fourth and final section has 3 1/2 dB of loss. This final section of the attenuator 213 comprises resistors 235 and 236 and a switch 237. By selecting the proper combination of on/off values for switches 224,229,233 and 237, the computer 110 may select from 0 to 56 dB of loss in 3 1/2 dB

20 increments. It should be understood that if a finer control of this attenuator is desired such that it could select attenuation in 1.75 dB increments, it is but a simple matter for one skilled in the art, in view of the above teachings, to add another section to the attenuator thereby providing this level of control.

    This signal from the programmable attenuator 213 is coupled to the low pass filter 215 which provides additional shaping to the transmit signal. Low pass filter 215 comprises a follower amplifier 238, and

25 associated circuitry comprising capacitors 239 and 240, and resistors 241 and 242. The output of filter 215 is coupled to a transmit audio output level conversion circuit, comprising amplifier 144, resistors 145, 146 and 147, and also capacitor 148, for connection to the audio output line 101. This output level conversion circuit provides an output impedance of 600 ohms for matching to the output line 101.

    Referring now to FIG. 5, there is shown a detail schematic for the programmable attenuator 313, the low

30 pass filter 315 and the amplifier 114 for the loudspeaker 112. The same basic components are used in implementing the programmable attenuator 313 and the programmable attenuator 213. Because of this and the detailed description given to attenuator 213, this attenuator 313 will not be described in similar detail.

    Follower amplifiers 326, 330 and 334 along with resistors 322, 323, 327, 328, 331, 332, 335 and 336, and also switches 324,329,333 and 337 combine in forming the four sections of the attenuator 313. As in

35 attenuator 213, a speech signal is attenuated 28 dB by section one, 14 dB by section two and 7 dB and 3 1/2 dB by sections three and four respectively.

    The signal from the programmable attenuator 313 is coupled to the low pass filter 315 which provides additional shaping to the receive signal. Low pass filter 315 comprises a follower amplifier 338, and associated circuitry including capacitors 339 and 340, and resistors 341 and 342. In amplifier 114, an

40 amplifier unit 149 and associated circuitry, variable resistor 150, resistors 151 and 152, and capacitors 153 and 154, provide gain for the output signal from low pass filter 315 before coupling this signal to the speaker 112 via a capacitor 155.

    With reference to FIG. 6, there is shown a general speakerphone circuit 600 for describing the two type of coupling, hybrid and acoustic, that most affect the operation of a speakerphone being employed in a

45 telephone connection. A hybrid 610 connects the transmit and receive paths of the speakerphone to a telephone line whose impedance may vary depending upon, for example, its length from a central office, as well as, for example, other hybrids in the connection. And the hybrid only provides a best case approximation to a perfect impedance match to this line. Thus a part of the signal on the transmit path to the hybrid returns over the receive path as hybrid coupling. With this limitation and the inevitable acoustic

50 coupling between a loudspeaker 611 and a microphone 612, transmit and receive loss controls 613 and 614 are inserted in the appropriate paths to avoid degenerative feedback or singing.

    In accordance with the invention, the computer controlled adaptive speakerphone 100 of FIG. 1 advantageously employs a process or program described herein with reference to a state diagram of FIG 7 and flow diagrams of FIGS. 8, 9 and 10 for improved performance. This process dynamically adjusts the

55 operational parameters of the speakerphone for the best possible performance in view of existing hybrid and acoustic coupling conditions.

    Referring now to FIG. 7, there is shown the state diagram depicting the possible states of the speakerphone 100. The speakerphone initializes in an Idle state 701. While in this state, the speakerphone

has a symmetrical path for entering into either a transmit state 702 or a receive state 703, according to which of these two has the stronger signal. If there is no transmit or receive speech while the speakerphone is in the idle state 701, the speakerphone remains in this state as indicated by a loop out of and back into this idle state. Generally, if speech is detected in the transmit or receive path, the speakerphone moves to the corresponding transmit or receive state. If the speakerphone has moved to the transmit state 702, for example, and transmit speech continues to be detected, the speakerphone then remains in this state. If the speakerphone detects receive speech having a stronger signal than the transmit speech, a receive break-in occurs and the speakerphone moves to the receive state 703. If transmit speech ceases and no receive speech is present, the speakerphone returns to the idle state 701. Operation of the speakerphone in the receive state 703 is essentially the reverse of its operation in the transmit state 702. Thus if there is receive speech following the speakerphone moving to the receive state 703, the speakerphone stays in this state. If transmit speech successfully interrupts, however, the speakerphone goes into the transmit state 702. And if there is no receive speech while the speakerphone is in the receive state 703 and no transmit speech to interrupt, the speakerphone returns to the idle state.

Referring next to FIG. 8, there is shown a flow chart illustrating in greater detail the operation of the speakerphone 100 in determining whether to remain in the idle state or move from the idle state to the transmit state or receive state. The process is entered at step 801 wherein the speakerphone is in the idle state. From this step, the process advances to the decision 802 where it determines whether the detected transmit signal is greater than the transmit noise by a certain threshold. If the detected transmit signal is greater than the transmit noise by the desired amount, the process proceeds to decision 803. At this decision, a determination is made as to whether the detected transmit signal exceeds the expected transmit signal by a certain threshold.

The expected transmit signal is that component of the transmit signal that is due to the receive signal coupling from the loudspeaker to the microphone. This signal will vary based on the receive speech signal, the amount of switched loss, and the acoustics of the room as determined during the acoustic calibration process. The expected transmit level is used to guard against false switching that can result from room echoes; therefore, the transmit level must exceed the expected transmit level by a certain threshold in order for the speakerphone to switch into the transmit state.

If the detected transmit signal does not exceed the expected transmit signal by the threshold, the process advances to decision 806. If the detected transmit signal exceeds the expected transmit signal by the threshold, however, the process advances to step 804 where a holdover timer is initialized prior to the speakerphone entering the transmit state. Once activated, this timer keeps the speakerphone in either the transmit state or the receive state over a period of time, approximately 1.2 seconds, when there is no speech in the then selected state. This allows a suitable period for bridging the gap between syllables, words and phrases that occur in normal speech. From step 804 the process advances to step 805 where the speakerphone enters the transmit state.

Referring once again to step 802, if the detected transmit signal is not greater than the transmit noise by a certain threshold, then the process advances to the decision 806. In this decision, and also in decision 807, the receive path is examined in the same manner as the transmit path in decisions 802 and 803. In decision 806, the detected received signal is examined to determine if it is greater than the receive noise by a certain threshold. If the detected receive signal is not greater than the receive noise by this threshold, the process returns to the step 801 and the speakerphone remains in the idle state. If the detected receive signal is greater than the receive noise by the desired amount, the process proceeds to decision 807. At this decision, a determination is made as to whether the detected receive signal exceeds the expected receive signal by a certain threshold.

The expected receive signal represents the amount of speech seen on the receive line that is due to transmit speech coupled through the hybrid. This signal is calculated on an ongoing basis by the speakerphone and depends on the hybrid average, the amount of switched loss, and the transmit speech signal. Since the transmit speech path is open to some extent while the speakerphone is in the idle state, this causes a certain amount of hybrid reflection to occur, which, in turn, causes a certain amount of the speech signal detected on the receive path to be due to actual background noise or speech in the room. This, in turn, is read as a certain expected level of receive speech. And the actual receive speech signal must surpass this expected level by the threshold in order for the speakerphone to determine with certainty that there is actually a far-end party talking.

If the detected receive signal does not exceed the expected receive signal by the threshold, the process returns to the step 801 and the speakerphone remains in the idle state. If the detected receive signal exceeds the expected receive signal by the threshold, however, the process advances to step 808 where the holdover timer is initialized. From step 808 the process advances to step 809 where the

speakerphone is directed to enter the receive state.

Referring next to FIG. 9, there is shown a flow chart illustrating in greater detail the operation of the speakerphone 100 in determining whether to remain in the transmit state or move from the transmit state to either the receive state or idle state. The process is entered at step 901 wherein the speakerphone has entered the transmit state. From this step, the process advances to the decision 902 where a determination is made as to whether the detected receive signal exceeds the expected receive signal by a certain threshold. If the detected receive signal does not exceed the expected receive signal by the threshold, the process advances to decision 907. If the detected receive signal exceeds the expected receive signal by the threshold, however, the process advances to step 903 where the detected received signal is examined to determine if it is greater than the receive noise by a certain threshold. If the detected receive signal is not greater than the receive noise by this threshold, the process advances to decision 907. If the detected receive signal is greater than the receive noise by the desired amount, the process proceeds to decision 904.

At decision 904, a determination is made as to whether the detected receive signal is greater than the detected transmit signal by a certain threshold. This decision is applicable when the near-end party and the far-end party are both speaking and the far-end party is attempting to break-in and change the state of the speakerphone. If the detected receive signal is not greater than the detected transmit signal by the threshold, the process proceeds to decision 907. If the detected receive signal is greater than the detected transmit signal by the threshold, however, the process proceeds to step 905 where the holdover timer is initialized for the receive state. From step 905, the process advances to step 906 where it causes the speakerphone to enter the receive state.

At decision 907, the process checks to see if the detected transmit signal is greater than the transmit noise by a certain threshold. If the detected transmit signal is greater than the transmit noise by the desired amount, the holdover timer is reinitialized at step 908, the process returns to step 901 and the speakerphone remains in the transmit state. Each time the holdover timer is reinitialized for a certain state, the speakerphone will remain minimally in that state for the period of the holdover timer, 1.2 seconds.

If at decision 907, the process finds that the detected transmit signal is less than the transmit noise by a certain threshold, i.e., no speech from the near-end party, the process advances to the decision 909 where it determines if the holdover timer has expired. If the holdover timer has not expired, the process returns to step 901 and the speakerphone remains in the transmit state. If the holdover timer has expired, the process advances to step 910 and the speakerphone returns to the idle state.

Referring next to FIG. 10, there is shown a flow chart illustrating in greater detail the operation of the speakerphone 100 in determining whether to remain in the receive state or move from the receive state to either the transmit state or idle state. The process is entered at step 1001 wherein the speakerphone has entered the receive state. From this step, the process advances to the decision 1002 where a determination is made as to whether the detected transmit signal exceeds the expected transmit signal by a certain threshold. If the detected transmit signal does not exceed the expected transmit signal by the threshold, the process advances to decision 1007. If the detected transmit signal exceeds the expected transmit signal by the threshold, however, the process proceeds to step 1003 where the detected transmit signal is examined to determine if it is greater than the transmit noise by a certain threshold. If the detected transmit signal is not greater than the transmit noise by this threshold, the process advances to decision 1007. If the detected transmit signal is greater than the transmit noise by the desired amount, the process proceeds to decision 1004.

At decision 1004, a determination is made as to whether the detected transmit signal is greater than the detected receive signal by a certain threshold. This decision is applicable when the far-end party and the near-end party are both speaking and the near-end party is attempting to break-in and change the state of the speakerphone. If the detected transmit signal is not greater than the detected receive signal by the threshold, the process proceeds to decision 1007. If the detected transmit signal is greater than the detected receive signal by the threshold, however, the process proceeds to step 1005 where the holdover timer is initialized for the transmit state. From step 1005, the process advances to step 1006 where it causes the speakerphone to enter the transmit state.

At decision 1007, the process checks to see if the detected receive signal is greater than the receive noise by a certain threshold. If the detected receive signal is greater than the receive noise by the desired amount, the holdover timer is reinitialized at step 1008, the process returns to step 1001 and the speakerphone remains in the receive state.

If at decision 1007, the process finds that the detected receive signal is less than the receive noise by a certain threshold, i. e., no speech from the far-end party, the process advances to the decision 1009 where it determines if the holdover timer has expired. If the holdover timer has not expired, the process returns to

10

step 1001 and the speakerphone remains in the receive state. If the holdover timer has expired, the process advances to step 1010 and the speakerphone returns to the idle state.

Referring now to FIG. 11, there is shown illustrative waveforms which provide an impulse and a composite characterization of an acoustic environment obtained during the acoustic calibration process performed by the speakerphone 100. A tone signal, generated between 300 Hz and 3.3 KHz in fifty equal logarithmically spaced frequency steps, is applied to the loudspeaker 112 of the speakerphone and the return echo for each tone measured by the microphone 111 and analyzed by the computer 110. Samples of the return echo for each tone signal generated are taken at 10 millisecond intervals for a retoal sampling period of 120 milliseconds.

The sample impulse responses shown in FIG. 11 are for the four frequencies, 300 Hz, 400 Hz, 500 Hz and 3.3 KHz. As illustrated in this figure, the 300 Hz response initially has a fairly high amplitude (A), but the energy quickly dissipates after the tone stops. In the 400 Hz response, its amplitude (A) is initially lower, however, the energy does not dissipate as rapidly as in the 300 Hz response. And the energy in the 500 Hz response dissipates even slower than the 300 Hz and the 400 Hz impulse responses.

A composite waveform is generated against each 300 Hz, 400 Hz and 500 Hz impulse response. This composite waveform represents an integrated response pattern of the impulse responses. The 300 Hz impulse response and the 300 Hz composite response are identical since this is the first measured response. The subsequent composite responses are modified based on the new information that comes in with each new impulse response. If that new information shows any ten millisecond time interval with a higher amplitude return than is then on the composite response for the corresponding time interval, the old information is replaced by the new information. If the new information has a lower amplitude return than that on the composite for that corresponding time interval, the old information is retained on the composite response. The 3.3 KHz frequency tone is the last of the 50 tones to be generated. The composite response after this tone represents, for each ten millisecond time interval, essentially the worst case acoustic coupling that may be encountered by the speakerphone during operation, independent of frequency.

This measure of the initial characterization of the room acoustic environment in which the speakerphone operates is used in a number of ways. The composite response is used for setting a switchguard threshold which insures that receive speech, if coming out of the loudspeaker is not falsely detected as transmit speech and returned to the far-end party.

The composite response is also used for determining the total amount of loop loss necessary for proper operation of the speakerphone. The amount of receive speech signal that is returned through the microphone from the loudspeaker is used as part of the equation which also includes the amount of hybrid return, the amount of loss inserted by the programmable attenuators and the gain setting of the volume control to determine the total amount of loop loss.

The composite response is further used in determining the expected transmit level. This expected transmit level is obtained from a convolution of the composite impulse response with the receive speech samples. The receive speech samples are available in real time for the immediately preceding 120 milliseconds with sample points at approximately 10 millisecond intervals. The value of the sample points occurring at each 10 millisecond interval in the receive response are convolved with the value of the sample points corresponding to the same 10 millisecond intervals in the composite response. In this convolution, the sampled values of the received speech response are, on a sample point by sample point basis, multiplied by the corresponding values of the sample points contained in the composite response. The resulting products are then summed together to obtain a single numerical value which represents the convolution of the immediately preceding 120 milliseconds of receive speech and 120 milliseconds of initial room characterization. This numerical value represents the amount of receive speech energy that is still in the room and will be detected by the microphone.

The following example illustrates how the convolution of the composite response with the received speech provides for more efficient operation of the speakerphone. If, by way of example, the near-end party begins talking and the speakerphone is in the receive state receiving speech from the far-end party, a certain amount of the signal coming out of the loudspeaker is coupled back into the microphone. The speakerphone has to determine whether the speech seen at the microphone is due solely to acoustic coupling, or whether it is due to the near-end talker. This determination is essential in deciding which state the speakerphone should be entering. To make this determination, the computer convolves the composite impulse response of the room with the receive speech signal to determine the level of speech seen at the microphone that is due to acoustic coupling. If the amount of signal at the microphone is greater than expected, then the computer knows that the near-end user is trying to interrupt and can permit a break-in; otherwise, the speakerphone will remain in the receive state.

When a speakerphone type device is operated in a near full or full duplex mode, the far-end party's

speech emanating from the loudspeaker is coupled back into the microphone and back through the telephone line to the far-end. Because of the proximity of the loudspeaker to the microphone, the speech level at the microphone resulting from speech at the loudspeaker is typically much greater than that produced by the near-end party. The result is a loud and reverberant return echo to the far-end. To alleviate this unpleasant side effect of near full or full duplex operation, an echo suppression process, which inserts loss in the transmit path as appropriate, is employed.

A diagram generally illustrating the insertion of echo suppression loss during near full or full duplex operation is shown in FIG 12. The speech signal in the receive path is measured by a measuring system 1210. Such a measuring system, by way of example, is available from high pass filter 312, envelope detector 314 and logarithmic amplifier 316 shown in FIG.1. The output of measuring system 1210 is passed through an acoustic coupling equation 1211 in order to include the effects of acoustic coupling on the signal to be seen at the microphone. The acoustic coupling equation could be as simple as a fast attack, slow decay analog circuit. In this implementation, the acoustic coupling equation is the composite room impulse response that is generated during the acoustic coupling phase of the calibration process. The output of the equation is the expected transmit signal level described earlier herein. The resulting signal is then used to provide a control signal for the modulation of the transmit path loss. An echo threshold detection circuit 1212 monitors the amplitude of the control signal from the acoustic coupling equation 1211. When the control signal exceeds a predetermined threshold (below which the return echo would not be objectionable to the far-end party) transmit loss which tracks the receive signal is inserted into the transmit path by the modulation circuit 1213.

By monitoring the transmit and receive speech signals, the process determines when the speech signal into the microphone is a result of acoustically coupled speech from the loudspeaker. While the speaker-phone is operating, the expected transmit signal level is also constantly monitored. This level is a direct indication of loudspeaker to microphone coupling and loop switched loss. This expected transmit signal level will tend to get larger as the speakerphone approaches full duplex operation. When this signal exceeds an echo threshold (below which the return echo would not be objectionable to the far-end party), additional loss is inserted into the transmit path. This echo suppression loss, when needed, tracks the receive speech envelope at a syllabic rate after a 1 to 5 millisecond delay.

Referring next to FIG. 13, there is shown a flow diagram illustrating the decision making process for the application of echo suppression loss. The process is entered at decision 1301 where the transmit signal level is compared with the expected transmit signal level plus a coupling threshold. If the expected transmit signal level plus the coupling threshold is less than the measured transmit signal, the process advances to step 1302 since receive speech is not present and echo suppression is therefore not necessary. If the expected transmit signal level plus the coupling threshold is greater than the measured transmit signal, the process advances to decision 1303 since the speakerphone is emanating speech from the loudspeaker that may need to be suppressed.

At decision 1303, a determination is made as to whether the loop switched loss is great enough to obviate the need for additional echo suppression loss. If loop switched loss is greater than the coupling threshold, the process advances to step 1304 since the switched loss will prevent objectionable echo to the far-end and echo suppression is not necessary. If loop switched loss is not great enough to provide sufficient echo reduction, however, the process advances to decision 1305.

At decision 1305, a determination is made as to whether the expected level of the transmit signal is greater than the loop switched loss plus an echo threshold. If so, the process advances to step 1306 since the return echo would not be objectionable to the far-end party and echo suppression is not necessary. If, however, the expected level of the transmit signal is less than the loop switched loss plus an echo threshold, echo suppression is necessary and the process advances to step 1307. The echo suppression is then inserted into the transmit path at step 1307 as follows: loss = expected transmit level - (loop switched loss - echo threshold).

Shown in FIG. 14 is a waveform illustrating how, in speakerphone 100, loss is inserted into the transmit path via programmable attenuator 213 in accordance with the echo suppression process.

Although a specific embodiment of the invention has been shown and described, it will be understood that it is but illustrative and that various modifications may be made therein without departing from the spirit and scope of the invention as defined in the appended claims.

Claims

1. A voice switching apparatus (100) for processing speech signals on a communication line (101,102),

including means for switching between a receive state for receiving speech signals from the communication line and a transmit state for transmitting speech signals over the communication line CHARACTERIZED IN INCLUDING,

a line adapting arrangement for determining the type of communication line to which the voice switching apparatus is connected, the line adapting arrangement comprising:

means (110, 200) for measuring a transmit speech signal provided by the apparatus to the communication line for transmission over the communication line;

means (110, 300) for measuring a receive speech signal from the communication line, the level of the receive speech signal being indicative of the return level of the transmit speech signal provided by the apparatus to the communication line; and

calibration means (110) operably responsive to both the transmit speech measuring means and the receive speech measuring means for adjusting threshold switching levels at which the apparatus switches between the receive state and the transmit state.

2. The line adapting arrangement as in claim 1 further CHARACTERIZED IN THAT the receive speech measuring means is operable only while the voice switching apparatus is in the transmit state.

3. The line adapting arrangement as in claim 1 further CHARACTERIZED IN INCLUDING variable switched loss means (213, 313) for alternately inserting loss in a receive path (310-316) for attenuating the speech signals received from the communication line and in a tansmit path (210-215) for attenuating the speech signals for transmission over the communication line;

the calibration means operably responsive to both the transmit speech measuring means and the receive speech measuring means for adjusting the level of attenuation inserted by the variable switched loss means into the transmit path and the receive path.

4. The line adapting arrangement as in claim 3 wherein the calibration means adjusts the level of attenuation of the variable loss means within a given range up to a predetermined incremental amount for each transition of the voice switching apparatus into the transmit state.

5. A method of determining the type of communication line to which a voice signal controller is connected, the voice signal controller being connectable to a communication line and switching between a receive state for receiving speech signals from the communication line and a transmit state for transmitting speech signals over the communication line, the method CHARACTERIZED IN COMPRISING the steps of:

measuring a transmit speech signal provided by the voice signal controller to the communication line for transmission over the communication line;

measuring a receive speech signal from the communication line, the level of the receive speech signal being indicative of the return level of the transmit speech signal provided by the controller to the communication line; and

adjusting threshold switching levels at which the controller switches between the receive state and the transmit state responsive to both the transmit speech signal measuring step and the receive speech signal measuring step.

6. The method of determining the type of communication line as in claim 5 further CHARACTERIZED IN THAT the receive speech measuring step is operable only while the voice signal controller is in the transmit state.

7. The method of determining the type of communication line as in claim 5 further CHARACTERIZED IN THAT comprising the steps of:

inserting loss alternately in a receive path for attenuating the speech signals received from the communication line and in a transmit path for attenuating the speech signals for transmission over the communication line; and

adjusting the level of attenuation inserted by the loss insertion step responsive to both the transmit speech signal measuring step and the receive speech signal measuring step.

8. The method of determining the type of communication line as in claim 7 further CHARACTERIZED IN THAT the attenuation level adjusting step adjusts the level of attenuation provided by the loss inserting step within a given range up to a predetermined incremental amount for each transition of the voice signal controller into the transmit state.

FIG.1

FIG.2

FIG.3

FIG.4

FIG.5

FIG.6

RX
LOSS CONTROL
600

611

610

H

HYBRID
COUPLING

613

614

TX
LOSS CONTROL

612

ACOUSTIC
COUPLING

FIG.7

TX
SPEECH   702

RX BREAK-IN

703   RX
SPEECH

TX

RX

TX BREAK-IN

NO TX
SPEECH   701

NO RX
SPEECH

TX
SPEECH

IDLE

RX
SPEECH

NO SPEECH
ON EITHER PATH

FIG.8

(IN)

IDLE — 801

802
NO ← IS $\frac{TX-S}{TX-N}$ >Th ?

YES

803
IS $\frac{TX-S}{TX-E}$ >Th ? — YES → INITIALIZE HOLDOVER TIMER — 804 → TX — 805

NO

806
NO ← IS $\frac{RX-S}{RX-N}$ >Th ?

YES

807
NO ← IS $\frac{RX-S}{RX-E}$ >Th ?

YES

808
INITIALIZE HOLDOVER TIMER

809
RX

KEY
TX-S=TRANSMIT SIGNAL
TX-N=TRANSMIT NOISE
RX-S=RECEIVE SIGNAL
RX-N=RECEIVE NOISE
TX-E=EXPECTED TRANSMIT
RX-E=EXPECTED RECEIVE
TH=THRESHOLD

FIG.9

TX ─ 901

IS $\dfrac{RX-S}{RX-E}$ >Th ? ─ 902
- NO (left)
- YES (down)

IS $\dfrac{RX-S}{RX-N}$ >Th ? ─ 903
- NO (left)
- YES (down)

IS $\dfrac{RX-S}{TX-S}$ >Th ? ─ 904
- YES → INITIALIZE HOLDOVER TIMER ─ 905
- NO (down)

INITIALIZE HOLDOVER TIMER → RX ─ 906

IS $\dfrac{TX-S}{TX-N}$ >Th ? ─ 907
- YES → REINITIALIZE HOLDOVER TIMER ─ 908
- NO (down)

HOLDOVER TIMER HAS EXPIRED ─ 909
- NO (left)
- YES (down)

IDLE ─ 910

KEY
TX-S=TRANSMIT SIGNAL
TX-N=TRANSMIT NOISE
RX-S=RECEIVE SIGNAL
RX-N=RECEIVE NOISE
TX-E=EXPECTED TRANSMIT
RX-E=EXPECTED RECEIVE
Th=THRESHOLD

FIG.10

```
                              ┌──────┐
                         ╭────┤  RX  │ ─── 1001
                         │    └──┬───┘
                         │       │
                         │       ▼
                    NO   ╱╲ ─── 1002
              ╭──────────╱ IS╲
              │          ╲ TX-S ╱
              │          ╲ ──── >Th ?╱
              │          ╲ TX-E ╱
              │           ╲╱
              │            │ YES
              │            ▼
              │       NO  ╱╲ ─── 1003
              ├──────────╱ IS╲
              │          ╲ TX-S ╱
              │          ╲ ──── >Th ?╱
              │          ╲ TX-N ╱
              │           ╲╱
              │            │ YES
              │            ▼                    ┌─────────────┐ ─── 1005
              │           ╱╲ ─── 1004      YES  │ INITIALIZE  │
              │          ╱ IS╲ ────────────────▶│  HOLDOVER   │
              │          ╲ TX-S ╱               │   TIMER     │
              │          ╲ ──── >Th ?╱          └──────┬──────┘
              │          ╲ RX-S ╱                      │
              │           ╲╱                           ▼
              │            │ NO                    ┌──────┐
   ┌──────────────┐        │                       │  TX  │
   │ REINITIALIZE │  YES  ╱╲ ─── 1007              └──────┘
   │   HOLDOVER   │◀─────╱ IS╲                      1006
   │    TIMER     │      ╲ RX-S ╱
   └──────────────┘      ╲ ──── >Th ?╱
     1008                ╲ RX-N ╱
              │           ╲╱
              │            │ NO
              │            ▼
              │           ╱╲ ─── 1009
        NO    │          ╱HOLDOVER╲
   ╭──────────┴─────────╱  TIMER HAS ╲
   │                    ╲  EXPIRED   ╱
   │                     ╲╱
   │                      │ YES ─── 1010
   │                      ▼
   │                  ┌──────┐
   │                  │ IDLE │
   │                  └──────┘
```

KEY

TX-S=TRANSMIT SIGNAL
TX-N=TRANSMIT NOISE
RX-S=RECEIVE SIGNAL
RX-N=RECEIVE NOISE
TX-E=EXPECTED TRANSMIT
RX-E=EXPECTED RECEIVE
Th=THRESHOLD

FIG.11

ACOUSTIC CALIBRATION

IMPULSE RESPONSE PATTERNS

INTEGRATED RESPONSE PATTERNS

FIG.12

RECEIVE
SPEECH
SIGNAL →

1210 — MEASURING
SYSTEM

1211 — ACOUSTIC
COUPLING
EQUATION

1212 — ECHO THRESHOLD
DETECTION

TRANSMIT SPEECH
SIGNAL IN →

1213 — TRANSMIT PATH
LOSS MODULATION

→ TRANSMIT SPEECH
SIGNAL OUT

FIG.13

```
                    (IN)
                      |
                      |      ┌ 1301                      ┌ 1302
                      v                          ┌──────────────────┐
                  /         \                    │     NO ECHO      │
                 /  TRANSMIT  \                  │   SUPRESSION     │
                /      <       \                 │     NEEDED       │
               <(EXPECTED TRANS + COUPLING)──────>│  (NO RECEIVE     │
                \  THRESHOLD)  /                  │ SPEECH PRESENT)  │
                 \           /                    └──────────────────┘
                  \         /
                      |
                      |      ┌ 1303                      ┌ 1304
                      v                          ┌──────────────────┐
                  /         \                    │     NO ECHO      │
                 /   LOOP     \                  │   SUPRESSION     │
                <SWITCHED LOSS < COUPLING>───────>│     NEEDED       │
                 \ THRESHOLD  /                   │ (SUFFICIENT      │
                  \         /                     │ SWITCHED LOSS)   │
                      |                           └──────────────────┘
                      |      ┌ 1305                      ┌ 1306
                      v                          ┌──────────────────┐
                  /         \                    │     NO ECHO      │
                 /  EXPECTED  \                  │   SUPRESSION     │
                <TRANSMIT > (SWITCHED LOSS)──────>│     NEEDED       │
                 \ + ECHO THRESHOLD)/             │ (INSUFFICIENT    │
                  \         /                     │  ECHO LEVEL)     │
                      |                           └──────────────────┘
                      |      ┌ 1307
                      v
          ┌─────────────────────────────┐
          │  INSERT ECHO SUPRESSION      │
          │                              │
          │ TRACKING LOSS = (EXPECTED    │
          │   TRANS)-(SWITCHED LOSS      │
          │     -ECHO THRESHOLD)         │
          └─────────────────────────────┘
                      |
                      v
                  ( EXIT )
```

FIG.14

THRESHOLD

COUPLED SPEECH

MODULATED LOSS

0